Data Management and Graduate Students: A One-Credit Course Case Study

Kyrille DeBose *Head, Veterinary Medicine Library & Liaison to Animal Sciences*

USAIN
UNITED STATES AGRICULTURAL INFORMATION NETWORK

VIRGINIA TECH

Background & Course Objectives

Background

Goal: Create a course that focuses on developing best practices for managing data across a wide spectrum of disciplines.

Course Objectives Included:

- Evaluate the importance and benefits of data management
- Develop appropriate documentation of data
- Investigate data sources and data access
- Articulate ethical issues surrounding data management
- Locate tools available to analyze and interpret data
- Apply a working knowledge of data security, storage, and backup strategies and how these relate to data sharing
- Compare methods for data publishing and discuss the impact of citing data sets
- Justify characteristics of preferred and acceptable file formats for data preservation and archiving
- Construct a data management plan

Course Schedule & Modules

Course Schedule

GRAD 5024 SPRING 2017 SCHEDULE

UNIT 1: Overview of Data Management
- Module 1.1 (1/17-1/22): Importance of Data Management (*due: 1/25*)
- Module 1.2 (1/23-1/29): Data Privacy & Regulations (*due: 2/1*)
- Module 1.3 (1/30-2/5): Data Organization (*due: 2/8*)

All assignments from UNIT 1 DUE Friday, February 10, 2017 at 11:55 p.m.

UNIT 2: Documenting, Locating, and Using Data Appropriately
- Module 2.1 (2/6-2/12): Documenting Data & Metadata (*due: 2/15*)
- Module 2.2 (2/13-2/19): Data Repositories & Other Sources of Data (*due: 2/22*)
- Module 2.3 (2/20-2/26): Ethical Use (and creation) of Data (*due: 3/1*)

All assignments from UNIT 2 DUE Friday, March 3, 2017 at 11:55 p.m.

UNIT 3: Analytical Tools for Data, Data Security, & Data Sharing
- Module 3.1 (3/13-3/19): Data Analysis & Visualization Tools (*due: 3/22*)
- Module 3.2(3/20/-3/26): Data Security & Storage (*due: 3/29*)
- Module 3.3 (3/27-4/2): Data Sharing (*due: 4/5*)

All assignments from UNIT 3 DUE Friday, April 14, 2017 at 11:55 p.m.

UNIT 4: Publishing & Archiving Data
- Module 4.1 (4/3-4/9): Publishing Data (*due: 4/13*)
- Module 4.2 (4/10-4/16): Citing Data & Datasets (*due: 4/19*)
- Module 4.3 (4/17-4/23): Data Preservation & Archiving (*due: 4/26*)

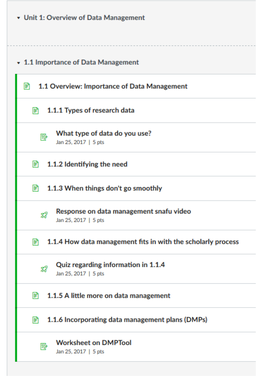
All assignments from UNIT 4 DUE Wednesday, May 3, 2017 at 11:55 p.m.

The Final Projects
There are three final projects for this course. They are:

- Data Management Plan
- ReadMe file of your Data Project
- Reflection on Publishing Options for your Dataset

All Final Projects are DUE Tuesday, May 9, 2017 at 11:55 p.m.

Overview of a Module



Module Objectives

1.2 Overview: Data Privacy & Regulations

Module Objectives

This week, we'll be covering the following topics:

- Describe the movement for Open Data
- Critique policies governing data privacy
- Examine the protocols and compliance training required for certain types of data collection
- Review different granting policies requirements for sharing data

Required Readings

- Website: PLoS Open Access
- PDF: OSTP Policy Memorandum
- PDF: Driving Innovation with Open Data
- Website: VT Starting Point (tool)

Assignments

- Response to readings (provided as an auto-graded quiz)
- Short narrative regarding data privacy (textbox entry)
- Worksheet on protocol and compliance training requirements
- Worksheet comparing data sharing policies

Due dates

- Complete section 1.2 by Wednesday, February 2, 2017 at 11:55pm
- Final due date for Unit 1 Friday, February 10, 2017 at 11:55 p.m

3.3 Overview: Data Analysis & Visualization Tools

This week's module is presented by our guest speakers, Chreston Miller (Data & Informatics Consultant, Engineering) and Michael Stamper (Data Visualization Designer and Digital Consultant for the Arts)

With your research data at hand, what can you do with it? In this module we will discuss the foundations of data analysis and visualization, plus showcase tools for discovering what your data is saying.

This week, we'll be covering the following topics:

- Examine data analysis tools
- Integrate a data analysis workflow
- Discuss the types of data visualization
- Select a data visualization tool to use

Required Readings

- Modified DataONE Module 9 powerpoint presentation
- Webpage: How to Tell a Story with Data
- Webpage: 7 Most Common Data Visualization Mistakes

Assignments

- Worksheet regarding data analysis tools and workflows
- Quiz comparing data visualizations
- Reflection on selecting data visualization options

Due dates

- Complete section 3.3 by Wednesday, April 5, 2017 at 11:55pm
- Final due date for Unit 3 Friday, April 14, 2017 at 11:55 p.m

New Course Approval Processes

Partnerships

Problem: The University Libraries are not able to offer a course directly through the student course catalog.

Solution: Build on a partnership with the Graduate School to offer a course through their programs.

Problem: Course leader did not have all of the skills to create applicable course content.

Solution: Utilized members of the Data Literacy & Consulting Working Group to create modules, activities, and assignments for specific areas based on their expertise.

Governance Approval

Tips:

- Identify timeline early on in the process & stick to it
- Locate someone who knows the system to review proposal
- Indicate how each learning outcome will be measured
- Listen to what others say, revise as needed, and have plenty of patience!

Students

Demographics

College of Agriculture and Life Sciences (6)
- Human Nutrition Foods & Exercise

College of Engineering (2)
- Mechanical Engineering
- Computer Science

Pamplin College of Business (1)

Virginia-Maryland College of Veterinary Medicine (2)
- Population Health Sciences

Student Feedback

Positives

- Appreciated in-depth feedback from instructor
- Promotion of an engaging environment between student and instructor

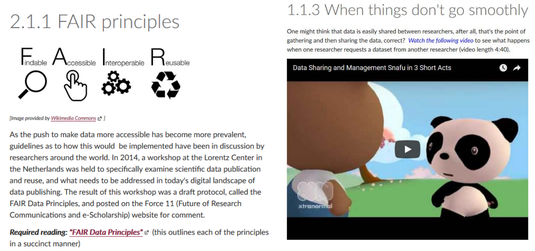
Negatives:

- Number of small assignments seen as "busywork"
- Too much work for a 1 credit course

Examples

Learning Objects

Modules incorporated a mix of text, images, videos, and audio clips.



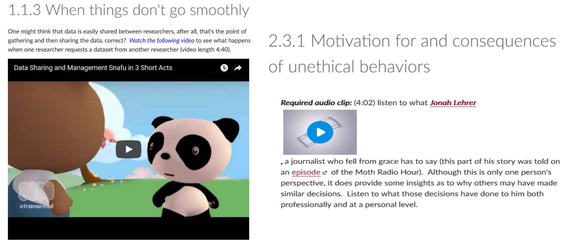
2.1.1 FAIR principles

Findable Accessible Interoperable Reusable

[Image provided by Wikimedia Commons]

As the push to make data more accessible has become more prevalent, guidelines as to how this would be implemented have been in discussion by researchers around the world. In 2014, a workshop at the Lorentz Center in the Netherlands was held to specifically examine scientific data publication and reuse, and what needs to be addressed in today's digital landscape of data publishing. The result of this workshop was a draft protocol, called the FAIR Data Principles, and posted on the Force 11 (Future of Research Communications and e-Scholarship) website for comment.

Required reading: *"FAIR Data Principles"* (this outlines each of the principles in a succinct manner)



1.1.3 When things don't go smoothly

One might think that data is easily shared between researchers, after all, that's the point of gathering and then sharing the data, correct? *Watch the following video* to see what happens when one researcher requests a dataset from another researcher (video length 4:40).

Data Sharing and Management Snafu in 3 Short Acts

2.3.1 Motivation for and consequences of unethical behaviors

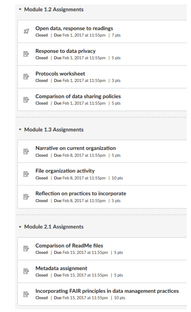
Required audio clip: (4:02) listen to what *Jonah Lehrer*

, a journalist who fell from grace has to say (this part of his story was told on an episode of the Moth Radio Hour). Although this is only one person's perspective, it does provide some insights as to why others may have made similar decisions. Listen to what those decisions have done to him both professionally and at a personal level.

Assignments & Activities

Each unit and grouping of final projects required 70% to pass:

- Each module comprised of 20 points
- Points were divided across 3-4 activities or assignments in each module
- All three final projects worth 20 points



Lessons Learned

Future Access to Content

No textbook for the course, but students requested an option to have access to the content once course completed.

- Created PDF versions of each unit
  - Included content (with full URLs) for the corresponding modules
  - Provided assignment/activities as supplementary files.
  - Included all images that were CC0 or in the public domain.

Changes Going Forward

1. Provide a mechanism to share student responses in assignments to the entire class to promote discussion.

2. Reduce required readings; create summaries with links to full source and offer as optional reading.

3. 75% of modules to have 2 assignments/activities; rest to have 1 practical-based assignment focusing on developing a specific skill set.